On Error Exponents in Quantum Hypothesis Testing

Tomohiro Ogawa [*], Masahito Hayashi [†]

Abstract

In the simple quantum hypothesis testing problem, upper bounds on the error probabilities are shown based on a key operator inequality between a density operator and its pinching. Concerning the error exponents, the upper bounds lead to a noncommutative analogue of the Hoeffding bound, which is identical with the classical counter part if the hypotheses, composed of two density operators, are mutually commutative. The upper bounds also provide a simple proof of the direct part of the quantum Stein's lemma.

Keywords

Hypothesis testing, Hoeffding bound, error exponent, quantum Stein's lemma, quantum relative entropy

1 Introduction

Quantum hypothesis testing is a fundamental problem in quantum information theory, because it is one of the most simple problems where the difficulty derived from noncommutativity of operators appears. It is also closely related to other topics in quantum information theory, as in classical information theory. Actually, its relation with quantum channel coding is discussed in [1] [2].

Let us outline briefly significant results in classical hypothesis testing for probability distributions $p^n(\cdot)$ versus $q^n(\cdot)$, where $p^n(\cdot)$ and $q^n(\cdot)$ are independently and identically distributed (i.i.d.) extensions of some probability distributions $p(\cdot)$ and $q(\cdot)$ on a finite set $\mathcal{X}$. In the classical case, the asymptotic behaviors of the first kind error probability α_n and the second kind error probability β_n for the optimal test were studied thoroughly as follows.

First, when α_n satisfies the constant constraint $\alpha_n \le \varepsilon$ ($\varepsilon > 0$), the error exponent of β_n for the optimal test is written asymptotically as

$$\lim_{n\to\infty} \frac{1}{n} \log \beta_n = -D(p||q) \tag{1}$$

[*]Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, 7–3–1 Hongo, Bunkyo-ku, Tokyo, 113–8656, Japan. (e-mail: ogawa@sr3.t.u-tokyo.ac.jp)

[†]Laboratory for Mathematical Neuroscience, Brain Science Institute, RIKEN, 2–1 Hirosawa, Wako, Saitama, 351–0198, Japan. (e-mail: masahito@brain.riken.go.jp)

for any ε (see *e.g.* [3], p.115), where $D(p||q)$ is the Kullback-Leibler divergence. The quantum analogue of (1) was established recently and called the quantum Stein's lemma [4] [5].

Next, when α_n satisfies the exponential constraint $\alpha_n \leq e^{-nr}$ $(r > 0)$, the error exponent of β_n for the optimal test is asymptotically determined by

$$\limsup_{n\to\infty} \frac{1}{n} \log \beta_n = - \min_{p' :\, D(p'||p) \leq r} D(p'||q) \tag{2}$$

$$= - \max_{0<s\leq 1} \frac{\Psi(s) - (1-s)r}{s}, \tag{3}$$

where the function $\Psi(s)$ is defined as

$$\Psi(s) \stackrel{\text{def}}{=} \sum_{x\in\mathcal{X}} p(x)^{1-s} q(x)^s. \tag{4}$$

Historically speaking, (2) and the test achieving it were shown in [6], followed by another expression (3) (see [7]), which we call the Hoeffding bound here. Concerning quantum fixed-length pure state source coding, an similar result is found in [8].

In this manuscript, a quantum analogue of the Hoeffding bound (3) (4) is introduced to derive a bound on the error exponent in quantum hypothesis testing. As a by-product of the process to derive the exponent, a simple proof of the quantum Stein's lemma is also given.

2 Definition and Main Results

Let $\mathcal{H}$ be a Hilbert space which represents a physical system in interest. We assume $\dim \mathcal{H} < \infty$ for mathematical simplicity. Let us denote the set of linear operators on $\mathcal{H}$ as $\mathcal{L}(\mathcal{H})$ and define the set of density operators on $\mathcal{H}$ by

$$\mathcal{S}(\mathcal{H}) \stackrel{\text{def}}{=} \{\rho \in \mathcal{L}(\mathcal{H}) \mid \rho = \rho^* \geq 0, \mathrm{Tr}[\rho] = 1\}. \tag{5}$$

We study the hypothesis testing problem for the null hypothesis $H_0 : \rho_n \stackrel{\text{def}}{=} \rho^{\otimes n} \in \mathcal{S}(\mathcal{H}^{\otimes n})$ versus the alternative hypothesis $H_1 : \sigma_n \stackrel{\text{def}}{=} \sigma^{\otimes n} \in \mathcal{S}(\mathcal{H}^{\otimes n})$, where $\rho^{\otimes n}$ and $\sigma^{\otimes n}$ are the nth tensor powers of arbitrarily given density operators ρ and σ in $\mathcal{S}(\mathcal{H})$.

The problem is to decide which hypothesis is true based on the data drawn from a quantum measurement, which is described by a positive operator valued measure (POVM) on $\mathcal{H}^{\otimes n}$, i.e. a resolution of identity $\sum_i M_{n,i} = I_n$ by nonnegative operators $M_n = \{M_{n,i}\}$ on $\mathcal{H}^{\otimes n}$. If a POVM consists of projections on $\mathcal{H}^{\otimes n}$, it is called a projection valued measure (PVM). In the hypothesis testing problem, however, it is sufficient to treat a two-valued POVM $\{M_0, M_1\}$, where the subscripts 0 and 1 indicate the acceptance of H_0 and H_1, respectively. Thus, an operator $A_n \in \mathcal{L}(\mathcal{H}^{\otimes n})$ satisfying inequalities $0 \leq A_n \leq I_n$ is called a test in the sequel, since A_n is identified with the POVM $\{A_n,\, I_n - A_n\}$. For a test A_n, the error probabilities of the first kind and the second kind are, respectively, defined by

$$\alpha_n(A_n) \stackrel{\text{def}}{=} \mathrm{Tr}[\rho_n(I_n - A_n)],$$
$$\beta_n(A_n) \stackrel{\text{def}}{=} \mathrm{Tr}[\sigma_n A_n].$$

Let us define the optimal value for $\beta_n(A_n)$ under the constant constraint on $\alpha_n(A_n)$:

$$\beta_n^*(\varepsilon) \stackrel{\text{def}}{=} \min\bigl\{\beta_n(A_n) \bigm| A_n : \text{test},\ \alpha_n(A_n) \le \varepsilon\bigr\}, \tag{6}$$

and let

$$D(\rho\|\sigma) \stackrel{\text{def}}{=} \mathrm{Tr}[\rho(\log\rho - \log\sigma)], \tag{7}$$

which is called the quantum relative entropy. Then we have the following theorem, which is one of the most essential theorems in quantum information theory.

Proposition 1 (The quantum Stein's lemma) *For $0 < \forall\varepsilon < 1$, it holds that*

$$\lim_{n\to\infty} \frac{1}{n} \log \beta_n^*(\varepsilon) = -D(\rho\|\sigma). \tag{8}$$

The first proof of (8) was composed of two inequalities, the direct part and the converse part. The direct part, concerned with existence of good tests, claims that

$$0 < \forall\varepsilon \le 1, \quad \limsup_{n\to\infty} \frac{1}{n} \log \beta_n^*(\varepsilon) \le -D(\rho\|\sigma), \tag{9}$$

and it was given by Hiai-Petz [4]. In this manuscript, the main focus is on the direct part, which is sometimes referred to as an equivalent form (see [5]):

$$\begin{aligned} &\exists\{A_n : \text{test}\}_{n=1}^{\infty} \quad \text{such that} \quad \lim_{n\to\infty} \alpha_n(A_n) = 0 \\ &\quad \text{and} \quad \limsup_{n\to\infty} \frac{1}{n} \log \beta_n(A_n) \le -D(\rho\|\sigma). \end{aligned} \tag{10}$$

On the other hand, the converse part, concerned with nonexistence of too good tests, asserts that

$$0 \le \forall\varepsilon < 1, \quad \liminf_{n\to\infty} \frac{1}{n} \log \beta_n^*(\varepsilon) \ge -D(\rho\|\sigma), \tag{11}$$

which was given by Ogawa-Nagaoka [5]. A direct proof of the equality (8) was also given by Hayashi [9] using the information spectrum approach in quantum setting [10], and a considerably simple proof of the converse part (11) was given in [11], recently.

In this manuscript, the asymptotic behavior of the error exponent $\frac{1}{n}\log\beta_n(A_n)$ under the exponential constraint $\alpha_n(A_n) \le e^{-nr}$ $(r > 0)$ is studied, and a noncommutative analogue of the Hoeffding bound [6] similar to (3) is given as follows.

Theorem 1 *For $\forall r > 0$, there exists a test A_n which satisfies*

$$\limsup_{n\to\infty} \frac{1}{n} \log \alpha_n(A_n) \le -r, \tag{12}$$

$$\limsup_{n\to\infty} \frac{1}{n} \log \beta_n(A_n) \le -\max_{0<s\le 1} \frac{\overline{\psi}(s) - (1-s)r}{s}, \tag{13}$$

where

$$\overline{\psi}(s) \stackrel{\rm def}{=} -\log {\rm Tr}\left[\rho\,\sigma^{\frac{s}{2}}\rho^{-s}\sigma^{\frac{s}{2}}\right]. \tag{14}$$

We will prove the theorem in Section 4. If ρ and σ are mutually commutative, $\overline{\psi}(s)$ is identical with the classical counterpart $\Psi(s)$ defined in (4), and (13) coincides with the Hoeffding bound (3), which is optimal in classical hypothesis testing.

This manuscript is organized as follows. In Section 3, upper bounds on the error probabilities are shown based on a key operator inequality [9]. Using the upper bounds, we will prove Theorem 1 in Section 4. In Section 5, the behavior of the function (14) is investigated, and a simple proof of the direct part of the quantum Stein's lemma (10) is given as a consequence of the upper bounds.

Two appendices are included for readers' convenience. Appendix A is devoted to the definition of the pinching map used effectively in Section 3. In Appendix B, the key operator inequality used in Section 3 is summarized briefly, along with another proof of it for readers' convenience.

3 Bounds on Error Probabilities

In the sequel, let $\mathcal{E}_{\sigma_n}(\rho_n)$ be the pinching defined in Appendix A and denote it as $\overline{\rho_n}$ for simplicity. Let $v(\sigma_n)$ be the number of eigenvalues of σ_n mutually different from others as defined in Appendix A. Then a key operator inequality [1] follows from Lemma 4 in Appendix B, which was originally appeared in [9]:

$$\rho_n \le v(\sigma_n)\,\overline{\rho_n}. \tag{15}$$

Note that the type counting lemma (see *e.g.* [12], Theorem 12.1.1) provides

$$v(\sigma_n) \le (n+1)^d, \tag{16}$$

where $d \stackrel{\rm def}{=} \dim \mathcal{H}$. Following [9], let us apply the operator monotonicity of the function $x \longmapsto -x^{-s}$ $(0 \le s \le 1)$ (see *e.g* [13]) to the key operator inequality (15) so that we have

$$\begin{aligned}\overline{\rho_n}^{\,-s} &\le v(\sigma_n)^s \rho_n^{-s} \\ &\le (n+1)^{sd}\rho_n^{-s}.\end{aligned} \tag{17}$$

Here, let us define the projection $\{X > 0\}$ for a Hermitian operator $X = \sum_i x_i E_i$ as

$$\{X > 0\} \stackrel{\rm def}{=} \sum_{i: x_i > 0} E_i, \tag{18}$$

where E_i is the projection onto the eigenspace corresponding to an eigenvalue x_i. With the above notation, we will focus on a test defined as

$$\overline{S}_n(a) \stackrel{\rm def}{=} \{\overline{\rho_n} - e^{na}\sigma_n > 0\}, \tag{19}$$

[1] Although the way to derive the operator inequality and the definition of $v(\sigma_n)$ are different from those of [9], it results in the same one as [9] in the case that both of ρ_n and σ_n are tensored states.

where a is a real parameter, and derive the upper bounds on the error probabilities for the test $\overline{S}_n(a)$ as follows.

Theorem 2

$$\alpha_n\left(\overline{S}_n(a)\right) \leq (n+1)^{sd}\, e^{-n\overline{\varphi}(a)}, \tag{20}$$

$$\beta_n\left(\overline{S}_n(a)\right) \leq (n+1)^{sd}\, e^{-n[\overline{\varphi}(a)+a]}, \tag{21}$$

where $\overline{\varphi}(a)$ *is defined by* $\overline{\psi}(s)$ *given in* (14) *as*

$$\overline{\varphi}(a) \stackrel{\text{def}}{=} \max_{0\leq s\leq 1}\left\{\overline{\psi}(s) - as\right\}. \tag{22}$$

Proof: The definition of $\overline{S}_n(a)$ and commutativity of operators $\overline{\rho_n}$ and σ_n lead to

$$\left(\overline{\rho_n}^{1-s} - e^{na(1-s)}\sigma_n^{1-s}\right)\overline{S}_n(a) \geq 0, \tag{23}$$

$$\left(\overline{\rho_n}^{s} - e^{nas}\sigma_n^{s}\right)\left(I_n - \overline{S}_n(a)\right) \leq 0 \tag{24}$$

for $\forall s \geq 0$. Note that $\overline{S}_n(a)$ also commutes with σ_n. Therefore, the inequality (24), with the property of the pinching (60) in Appendix A, provides

$$\begin{aligned}
\alpha_n\left(\overline{S}_n(a)\right) &= \mathrm{Tr}\left[\rho_n\left(I_n - \overline{S}_n(a)\right)\right] \\
&= \mathrm{Tr}\left[\overline{\rho_n}\left(I_n - \overline{S}_n(a)\right)\right] \\
&= \mathrm{Tr}\left[\overline{\rho_n}^{1-s}\overline{\rho_n}^{s}\left(I_n - \overline{S}_n(a)\right)\right] \\
&\leq e^{nas}\mathrm{Tr}\left[\overline{\rho_n}^{1-s}\sigma_n^{s}\left(I_n - \overline{S}_n(a)\right)\right] \\
&\leq e^{nas}\mathrm{Tr}\left[\overline{\rho_n}^{1-s}\sigma_n^{s}\right].
\end{aligned} \tag{25}$$

In the same way, (23) yields

$$\begin{aligned}
\beta_n\left(\overline{S}_n(a)\right) &= \mathrm{Tr}\left[\sigma_n\overline{S}_n(a)\right] \\
&= \mathrm{Tr}\left[\sigma_n^{s}\sigma_n^{1-s}\overline{S}_n(a)\right] \\
&\leq e^{-na(1-s)}\mathrm{Tr}\left[\sigma_n^{s}\overline{\rho_n}^{1-s}\overline{S}_n(a)\right] \\
&\leq e^{-na}e^{nas}\mathrm{Tr}\left[\overline{\rho_n}^{1-s}\sigma_n^{s}\right].
\end{aligned} \tag{26}$$

Here, it follows from the property (60) and (17) that

$$\begin{aligned}
\mathrm{Tr}\left[\overline{\rho_n}^{1-s}\sigma_n^{s}\right] &= \mathrm{Tr}\left[\overline{\rho_n}\,\sigma_n^{\frac{s}{2}}\,\overline{\rho_n}^{-s}\sigma_n^{\frac{s}{2}}\right] \\
&= \mathrm{Tr}\left[\rho_n\sigma_n^{\frac{s}{2}}\,\overline{\rho_n}^{-s}\sigma_n^{\frac{s}{2}}\right] \\
&\leq (n+1)^{sd}\,\mathrm{Tr}\left[\rho_n\sigma_n^{\frac{s}{2}}\,\rho_n^{-s}\sigma_n^{\frac{s}{2}}\right] \\
&= (n+1)^{sd}\,\left(\mathrm{Tr}\left[\rho\,\sigma^{\frac{s}{2}}\rho^{-s}\sigma^{\frac{s}{2}}\right]\right)^{n} \\
&= (n+1)^{sd}\,e^{-n\overline{\psi}(s)}
\end{aligned} \tag{27}$$

for $0 \leq \forall s \leq 1$. Combining (25) (26) (27), we have

$$\alpha_n\left(\overline{S}_n(a)\right) \leq (n+1)^{sd}\, e^{-n\left[\overline{\psi}(s)-as\right]}, \tag{28}$$

$$\beta_n\left(\overline{S}_n(a)\right) \leq (n+1)^{sd}\, e^{-n\left[\overline{\psi}(s)-as+a\right]}, \tag{29}$$

which lead to (20) (21) by taking the maximum in the exponents. ■

4 Proof of Theorem 1

In this section, we will prove Theorem 1 after preparing two lemmas, where the behavior of $\overline{\varphi}(a)$ in the error exponents (20) (21) is investigated.

Lemma 1 *$\overline{\varphi}(a)$ is convex and monotonically nonincreasing.*

Proof: The assertion immediately follows from the definition of $\overline{\varphi}(a)$. Actually, we have for $0 \leq \forall t \leq 1$

$$\begin{aligned}
\overline{\varphi}(ta+(1-t)b) &= \max_{0\leq s\leq 1}\left\{\overline{\psi}(s)-(ta+(1-t)b)s\right\} \\
&\leq t\max_{0\leq s\leq 1}\left\{\overline{\psi}(s)-as\right\}+(1-t)\max_{0\leq s\leq 1}\left\{\overline{\psi}(s)-bs\right\} \\
&= t\overline{\varphi}(a)+(1-t)\overline{\varphi}(b).
\end{aligned} \tag{30}$$

Next, let $a \leq b$ and $s_b \stackrel{\text{def}}{=} \arg\max_{0\leq s\leq 1}\left\{\overline{\psi}(s)-bs\right\}$. Then we have

$$\begin{aligned}
\overline{\varphi}(b) &= \overline{\psi}(s_b)-bs_b \\
&\leq \overline{\psi}(s_b)-as_b \\
&\leq \max_{0\leq s\leq 1}\left\{\overline{\psi}(s)-as\right\} \\
&= \overline{\varphi}(a).
\end{aligned} \tag{31}$$

■

Lemma 2 *$\overline{\varphi}(a)$ ranges from 0 to infinity.*

Proof: Since we can calculate the derivative of $\overline{\psi}(s)$ explicitly, $\overline{\psi}(s)$ is continuous and differentiable. Therefore, it follows from the mean value theorem that for $s > 0$ there exists $0 \leq t \leq s$ such that

$$\overline{\psi}'(t) = \frac{\overline{\psi}(s)-\overline{\psi}(0)}{s-0}. \tag{32}$$

Let $a \geq \max_{0\leq t\leq 1}\overline{\psi}'(t)$, then we have

$$a \geq \frac{\overline{\psi}(s)-\overline{\psi}(0)}{s-0}, \tag{33}$$

and hence

$$\overline{\psi}(0) \geq \overline{\psi}(s) - as, \tag{34}$$

which yields

$$0 = \overline{\psi}(0) = \max_{0 \leq s \leq 1} \left\{ \overline{\psi}(s) - as \right\} = \overline{\varphi}(a). \tag{35}$$

On the other hand, it is obvious that

$$\lim_{a \longrightarrow -\infty} \overline{\varphi}(a) = \infty. \tag{36}$$

Since $\overline{\varphi}(a)$ is continuous, which follows from convexity by Lemma 1, the assertion follows from (35) and (36). ■

Combined with the above lemma, Theorem 2 leads to Theorem 1 as follows.

Proof of Theorem 1: For $\forall r > 0$, there exists $a_r \in \mathbb{R}$ such that $r = \overline{\varphi}(a_r)$ from Lemma 2. Let $\overline{u}(r) \stackrel{\text{def}}{=} \overline{\varphi}(a_r) + a_r$, then it follows from Theorem 2 that

$$\limsup_{n \to \infty} \frac{1}{n} \log \alpha_n \left(\overline{S}_n(a_r) \right) \leq -r, \tag{37}$$

$$\limsup_{n \to \infty} \frac{1}{n} \log \beta_n \left(\overline{S}_n(a_r) \right) \leq -\overline{u}(r). \tag{38}$$

Therefore, it suffices to show that

$$\overline{u}(r) = \max_{0 < s \leq 1} \frac{\overline{\psi}(s) - (1-s)r}{s}. \tag{39}$$

For $0 \leq \forall s \leq 1$, we have from the definition of $\overline{\varphi}(a)$

$$r = \overline{\varphi}(a_r) \geq \overline{\psi}(s) - a_r s, \tag{40}$$

and there exists a number s_0 $(0 < s_0 \leq 1)$ achieving the equality since $r = \overline{\varphi}(a_r) > 0$. On the other hand, the definitions of $u(r)$ and a_r lead to

$$\overline{u}(r) = \overline{\varphi}(a_r) + a_r = r + a_r. \tag{41}$$

Eliminating a_r from (40) and (41), we have

$$\overline{u}(r) \geq \frac{\overline{\psi}(s) - (1-s)r}{s}, \tag{42}$$

and s_0 achieves the equality in (42) as well. Thus, we have shown (39), and Theorem 1 has been proved. ■

5 Graphs of $\overline{\psi}(s)$ and $\overline{\varphi}(a)$

In this section, we will investigate the graphs of $\overline{\psi}(s)$ and $\overline{\varphi}(a)$. To this end, let us define

$$\psi(s) \stackrel{\text{def}}{=} -\log \mathrm{Tr}\left[\rho^{1-s}\sigma^s\right], \tag{43}$$

$$\varphi(a) \stackrel{\text{def}}{=} \max_{0\le s\le 1}\left\{\psi(s) - as\right\}. \tag{44}$$

Then we have the following lemma.

Lemma 3

$$\overline{\psi}(s) \le \psi(s) \quad (0 \le \forall s \le 1), \tag{45}$$

$$\overline{\varphi}(a) \le \varphi(a) \quad (\forall a \in \mathbb{R}). \tag{46}$$

Proof: Let us apply the monotonicity property of the quantum quasi-entropy [14] [15] to $\mathrm{Tr}\left[\rho^{1-s}\sigma^s\right]$ $(0 \le s \le 1)$ [2] so that we have

$$\begin{aligned} e^{-n\psi(s)} &= \left(\mathrm{Tr}\left[\rho^{1-s}\sigma^s\right]\right)^n \\ &= \mathrm{Tr}\left[\rho_n^{1-s}\sigma_n^s\right] \\ &\le \mathrm{Tr}\left[\overline{\rho_n}^{1-s}\sigma_n^s\right] \\ &\le (n+1)^{sd}\, e^{-n\overline{\psi}(s)}, \end{aligned} \tag{47}$$

where we used (27) in the last inequality. Thus, we obtain

$$\overline{\psi}(s) \le \psi(s) + \frac{sd}{n}\log(n+1) \tag{48}$$

for any positive number n, and we have (45) by letting n go to infinity. Now (46) is obvious from the definition of $\overline{\varphi}(a)$. Actually, let $s_a \stackrel{\text{def}}{=} \arg\max_{0\le s\le 1}\left\{\overline{\psi}(s) - as\right\}$, then we have

$$\begin{aligned} \overline{\varphi}(a) &= \overline{\psi}(s_a) - as_a \\ &\le \psi(s_a) - as_a \\ &\le \max_{0\le s\le 1}\left\{\psi(s) - as\right\} \\ &= \varphi(a). \end{aligned} \tag{49}$$

■

Following [5], we can easily draw the graphs of $\psi(s)$ and $\varphi(a)$ (see Figure 1 and 2) by calculating the derivatives

$$\psi'(s) = e^{\psi(s)}\mathrm{Tr}\left[\rho^{1-s}\sigma^s\left(\log\rho - \log\sigma\right)\right], \tag{50}$$

$$\begin{aligned} \psi''(s) &= -e^{\psi(s)}\mathrm{Tr}\left[\rho^{1-s}A\sigma^s A\right] \\ &= -e^{\psi(s)}\mathrm{Tr}\left[\left(\rho^{\frac{1-s}{2}}A\sigma^{\frac{s}{2}}\right)\left(\rho^{\frac{1-s}{2}}A\sigma^{\frac{s}{2}}\right)^*\right] \\ &< 0, \end{aligned} \tag{51}$$

[2] A comprehensible explanation of the monotonicity property is found in [5].

where we put

$$A \stackrel{\text{def}}{=} \log\rho - \log\sigma - \psi'(s). \tag{52}$$

Especially, note that $\psi(0) = 0$ and $\psi'(0) = D(\rho||\sigma)$.

On the other hand, we can not know a lot concerning the graphs of $\overline{\psi}(s)$ and $\overline{\varphi}(a)$, except that we have $\overline{\psi}(0) = 0$ and $\overline{\psi}'(0) = D(\rho\|\sigma)$. Considering lemmas from 1 to 3, however, we can show the graphs of $\overline{\psi}(s)$ and $\overline{\varphi}(a)$ roughly as Figure 1 and 2. Here, it should be pointed out that

$$\overline{\varphi}(a) > 0 \quad \text{for} \quad \forall a < D(\rho||\sigma), \tag{53}$$

which leads to the following theorem combined with Theorem 2.

Theorem 3 *For $\forall a < D(\rho||\sigma)$, we have*

$$\lim_{n\to\infty} \alpha_n\left(\overline{S}_n(a)\right) = 0, \tag{54}$$

$$\limsup_{n\to\infty} \frac{1}{n}\log\beta_n\left(\overline{S}_n(a)\right) \le -a. \tag{55}$$

Since $a < D(\rho||\sigma)$ can be arbitrarily near $D(\rho||\sigma)$, we have shown the direct part of the quantum Stein's lemma (10).

6 Concluding Remarks

We have shown upper bounds on the error probabilities of the first and the second kind, based on a key operator inequality satisfied by a density operator and its pinching. The upper bounds are regarded as a noncommutative analogue of the Hoeffding bound [6], which is the optimal bound in the classical hypothesis testing, and the upper bounds provide a simple proof of the direct part of the quantum Stein's lemma. Compared with [9], the proof is considerably simple and leads to the exponential convergence of the error probability of the first kind.

The error exponents derived here do not seem to be natural, since $\overline{\psi}(s)$ lacks symmetry between ρ and σ that the original hypothesis testing problem has. One may introduce the following quantity as a substitute for $\overline{\psi}(s)$ to keep the symmetry:

$$\max\Big\{-\log\mathrm{Tr}\left[\rho\,\sigma^{\frac{s}{2}}\rho^{-s}\sigma^{\frac{s}{2}}\right], -\log\mathrm{Tr}\left[\sigma\,\rho^{\frac{s}{2}}\sigma^{-s}\rho^{\frac{s}{2}}\right]\Big\},$$

and Theorem 1 still holds with the above quantity. On the other hand, $\psi(s)$ and $\varphi(a)$ defined in (43) (44) seem to be probable functions for the optimal rate function in quantum hypothesis testing, and the following inequalities are expected to hold

$$\limsup_{n\to\infty} \frac{1}{n}\log\alpha_n\left(S_n(a)\right) \le -\varphi(a), \tag{56}$$

$$\limsup_{n\to\infty} \frac{1}{n}\log\beta_n\left(S_n(a)\right) \le -\big(\varphi(a) + a\big), \tag{57}$$

where

$$S_n(a) \stackrel{\rm def}{=} \{\rho_n - e^{na}\sigma_n > 0\}\,. \tag{58}$$

The question of whether the inequalities hold or not seems to be difficult, however, and is left open.

Appendices

A Definition of the Pinching

In this appendix, we summarize the definition of the pinching and some of its properties. Given an operator $A \in \mathcal{L}(\mathcal{H})$, let $A = \sum_{i=1}^{v(A)} a_i E_i$ be its spectral decomposition, where $v(A)$ is the number of eigenvalues of A mutually different from others, and each E_i is the projection corresponding to an eigenvalue a_i. The following map defined by using the PVM $E = \{E_i\}_{i=1}^{v(A)}$ is called the pinching:

$$\mathcal{E}_A : B \in \mathcal{L}(\mathcal{H}) \longmapsto \mathcal{E}_A(B) \stackrel{\rm def}{=} \sum_{i=1}^{v(A)} E_i B E_i \in \mathcal{L}(\mathcal{H}). \tag{59}$$

The operator $\mathcal{E}_A(B)$ is also called the pinching when no confusion is likely to arise, and it is sometimes denoted as $\mathcal{E}_E(B)$. It should be noted here that $\mathcal{E}_A(B)$ commutes with A and we have

$$\mathrm{Tr}[BC] = \mathrm{Tr}\left[\mathcal{E}_A(B)C\right] \tag{60}$$

for any operator $C \in \mathcal{L}(\mathcal{H})$ commuting with A.

B Key Operator Inequality

The following lemma was appeared in [9], and played an important role in this manuscript.

Lemma 4 (Hayashi [9]) *Given a PVM $M = \{M_i\}_{i=1}^{v(M)}$ on $\mathcal{H}$, we have for $\forall \rho \in \mathcal{S}(\mathcal{H})$*

$$\rho \le v(M)\mathcal{E}_M(\rho), \tag{61}$$

where $\mathcal{E}_M(\rho)$ is the pinching defined in Appendix A.

We show another proof here for readers' convenience by using the following operator convexity.

Lemma 5 *Given a nonnegative operator $A \in \mathcal{L}(\mathcal{H})$, the following map is operator convex.*

$$f_A : X \in \mathcal{L}(\mathcal{H}) \longmapsto X^*AX \in \mathcal{L}(\mathcal{H}). \tag{62}$$

In other words, we have

$$f_A(tX + (1-t)Y) \leq tf_A(X) + (1-t)f_A(Y) \tag{63}$$

for $\forall X, Y \in \mathcal{L}(\mathcal{H})$ and $0 \leq \forall t \leq 1$.

Proof: The assertion is shown by a direct calculation as follows

$$\begin{aligned}
&tf_A(X) + (1-t)f_A(Y) - f_A(tX + (1-t)Y) \\
&= tX^*AX + (1-t)Y^*AY - [tX + (1-t)Y]^*A\,[tX + (1-t)Y] \\
&= t(1-t)[X^*AX - X^*AY - Y^*AX + Y^*AY] \\
&= t(1-t)(X-Y)^*A\,(X-Y) \\
&\geq 0.
\end{aligned} \tag{64}$$

■

Now Lemma 4 is verified by using Lemma 5 as follows

$$\begin{aligned}
\frac{1}{v(M)^2}\rho &= \left(\frac{1}{v(M)}\sum_{i=1}^{v(M)} M_i\right)\rho\left(\frac{1}{v(M)}\sum_{i=1}^{v(M)} M_i\right) \\
&\leq \frac{1}{v(M)}\sum_{i=1}^{v(M)} M_i\rho M_i \\
&= \frac{1}{v(M)}\mathcal{E}_M(\rho).
\end{aligned} \tag{65}$$

Acknowledgment

The authors are grateful to Prof. Hiroshi Nagaoka. He encouraged them to show a simple proof of the direct part of the quantum Stein's lemma, pointing out that the proof leads to Hiai-Petz's theorem.

This research was partially supported by the Ministry of Education, Culture, Sports, Science, and Technology Grant-in-Aid for Encouragement of Young Scientists, 13750058, 2001.

References

[1] T. Ogawa and H. Nagaoka, "A new proof of the channel coding theorem via hypothesis testing in quantum information theory," to appear in ISIT2002.

[2] M. Hayashi and H. Nagaoka, "A general formula for the classical capacity of a general quantum channel," to appear in ISIT2002.

[3] R. E. Blahut, *Principles and Practice of Information Theory*, Addison-Wesley, Massachusetts, 1991.

[4] F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," *Commun. Math. Phys.*, vol. 143, pp. 99–114, 1991.

[5] T. Ogawa and H. Nagaoka, "Strong converse and Stein's lemma in quantum hypothesis testing," *IEEE Trans. Inform. Theory*, vol. IT-46, pp. 2428–2433, 2000.

[6] W. Hoeffding, "On probabilities of large deviations," Proceedings of Symposium "the Fifth Berkeley Symposium on Mathematical Statistics and Probability," pp. 203–219, Berkeley, University of California Press, 1965.

[7] R. E. Blahut, "Hypothesis testing and information theory," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 405–417, 1974.

[8] M. Hayashi, "Exponents of quantum fixed-length pure state source coding," LANL e-print quant-ph/0202002, 2002.

[9] M. Hayashi, "Optimal sequence of POVMs in the sense of Stein's lemma in quantum hypothesis testing," LANL e-print quant-ph/0107004, 2001.

[10] H. Nagaoka, "On asymptotic theory of quantum hypothesis testing," Proceedings of Symposium "Statistical inference theory and its information theoretical aspect," pp. 49–52, 1998, (In Japanese).

[11] H. Nagaoka, "Strong converse theorems in quantum information theory," Proceedings of Symposium "ERATO Workshop on Quantum Information Science 2001," p. 33, 2001.

[12] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.

[13] R. Bhatia, *Matrix Analysis*, Springer, New York, 1997.

[14] D. Petz, "Quasi-entropies for states of a von Neumann algebra," *Publ. RIMS, Kyoto Univ.*, pp. 787–800, 1985.

[15] D. Petz, "Quasi-entropies for finite quantum systems," *Rep. Math. Phys.*, vol. 23, pp. 57–65, 1986.

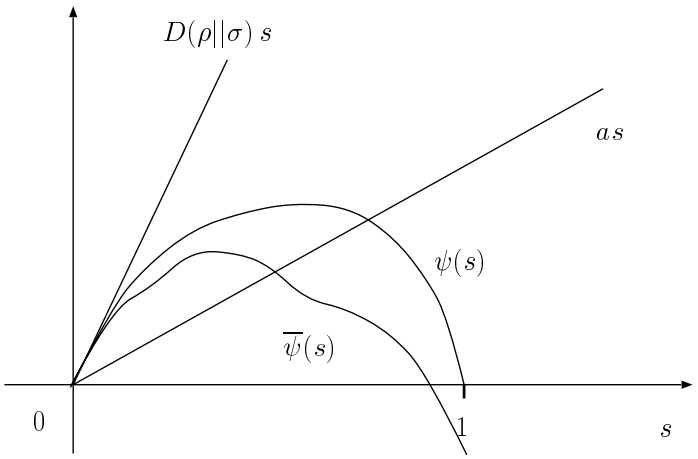

Figure 1: The graph of $\overline{\psi}(s)$

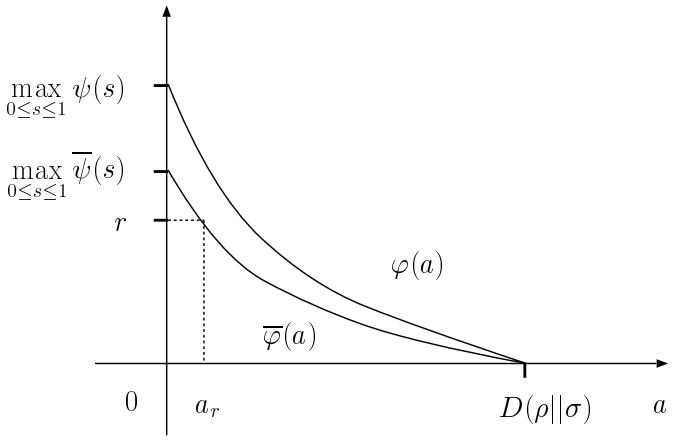

Figure 2: The graph of $\overline{\varphi}(a)$